



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.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



The income gradient in COVID-19 mortality and hospitalisation: An observational study with social security administrative records in Mexico

Eva O. Arceo-Gomez,^a Raymundo M. Campos-Vazquez,^{b*} Gerardo Esquivel,^c Eduardo Alcaraz,^d Luis A. Martinez,^d and Norma G. Lopez,^d

^aUniversidad Iberoamericana, Mexico City, Mexico

^bCentro de Estudios Económicos, El Colegio de México, Carretera Picacho Ajusco 20, Col. Ampliación Fuentes del Pedregal, Mexico City 14110, Mexico

^cBanco de México, Mexico City, Mexico

^dInstituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico

Summary

Background The COVID-19 pandemic revealed large structural inequalities that led to disparities in health outcomes related to socioeconomic status. So far, most of the evidence is based on aggregated data or simulations with individual data, which point to various possible mechanisms behind the association. To date, there have been no studies regarding an income gradient in COVID-19 mortality based on individual-level data and adjusting for comorbidities or access to healthcare.

Methods In this paper, we use linked employee-patient data for patients tested for COVID-19 at the Mexican Institute of Social Security. We estimate the association of the probability of dying with income centiles, using a probit estimation and adjusting for COVID-19 diagnosis, sociodemographic variables, and comorbidities.

Findings After controlling for all these variables, we find that persons in the lowest income decile still had a probability of dying from COVID-19 five times greater than those at the top decile.

Interpretation Our results imply that the association between income and COVID outcomes is not explained by the prevalence of comorbidities or by a lack of access to healthcare among the low-income population.

Funding This study was not supported by any external funding source.

Copyright © 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Keywords: COVID-19; Mortality; Hospitalisation; Comorbidities; Income gradient; Mexico

Introduction

As with many other diseases, the SARS-COV-2 virus interacts with structural inequalities to generate disparities in COVID-19 outcomes. Evidence points to socioeconomic characteristics and racial or ethnic origin as social determinants of these outcomes.^{1–3} The most commonly cited reasons for these disparities are pre-existing health conditions, higher risk of exposure to the virus, and inequalities in access to healthcare and treatment,^{3–7} leading some to describe the pandemic as a syndemic.^{8–10} However, when it comes to evidence for the association of poverty or income with COVID-19 outcomes, analyses are based mostly on socioeconomic or mortality data aggregated on a geographical level,

varying from small statistical areas^{11,12} or municipalities^{13,14} to countries.¹⁵ Other studies use limited individual-level samples, drawn from university health systems,⁶ hospitals,¹⁶ or state- or city-level healthcare systems,^{12,17} that are linked to sociodemographic data aggregated on a geographical level.

Two notable exceptions in the literature analyse individual income data^{18,19} by linking individual data from mortality records to income and other sociodemographic characteristics from administrative records in Sweden and Belgium. An unresolved issue in these studies is whether comorbidities are a confounder in the income gradient in COVID-19 mortality. In addition, both Sweden and Belgium are high-income countries with relatively low inequality levels, whose experience may not apply to low- and middle-income countries.

*Corresponding author.

E-mail address: rmcampos@colmex.mx (R.M. Campos-Vazquez).

The Lancet Regional Health - Americas

2022;6: 100115

Published online 10

November 2021

<https://doi.org/10.1016/j.lana.2021.100115>

100115

Research in context

Evidence before this study

Previous studies have found a negative association between income and COVID-19 mortality. Most of this research used data aggregated at a geographical level and was thus unable to estimate a direct association between individual income and COVID-19 outcomes. These studies have also not resolved whether the higher mortality among people with lower incomes is due to a higher infection rate, a higher incidence of comorbidities which aggravate COVID-19, or lesser access to healthcare. There are only two studies, one for Sweden and one for Belgium, that use individual-level data. In these studies, the authors linked mortality records to administrative records containing information on income and other sociodemographic variables. Both studies found a negative association between income and COVID-19 mortality. However, both also left open the question of whether the association between individual income and mortality is explained by the incidence of comorbidities or a higher infection rate among people with lower incomes.

Added value of this study

Our contributions to the literature are twofold. First, to our knowledge, we are the first to provide evidence addressing the income gradient in COVID-19 mortality using individual level data in a non-advanced and highly unequal economy such as Mexico. Individual-level information on income allows us to capture greater disparities than studies that use aggregated socioeconomic status (SES) measures. Second, we have linked two types of administrative records from the largest healthcare system in Mexico, the Instituto Mexicano del Seguro Social (IMSS): employment records for all IMSS-affiliated workers, including daily earnings, and medical records of patients tested for SARS-CoV-2 in IMSS facilities, which record comorbidities and COVID-19 outcomes. This unique dataset allows us to estimate the income gradient in COVID-19 outcomes, adjusting for the presence of comorbidities while exploiting greater dispersion in economic variables than suggested by aggregated data. Our results point to greater health inequality outcomes in the Mexican case than previously found in the literature: persons in the lowest earnings decile had a probability of dying from COVID-19 five times greater than those at the top decile, even after adjusting for comorbidities. We also provide evidence that people with lower income are not more likely to be tested for SARS-CoV-2, and they do not test positive at a higher rate. This evidence suggests that results are not being driven by higher infection rates among people with lower income.

Implication of all the available evidence

The COVID-19 pandemic has highlighted the disparities in health outcomes among different populations around the world. There is now ample evidence that the hardest hit subpopulations have been those with

the least economic resources, including racial and ethnic minorities. In this study we provide evidence that low-income workers in Mexico suffered from higher hospitalisation and fatality rates than their counterparts with higher wages, even though they did not necessarily have higher rates of testing or testing positive for SARS-CoV-2. Since all individuals in the sample had access to healthcare, there are likely other factors at play, and there is an opportunity to reduce this disparity by putting more emphasis on policies directed at people with lower incomes, such as emphasizing the need for timely diagnosis and medical attention.

In Mexico, the first confirmed cases of COVID-19 came at the end of February 2020. From then on, the pandemic evolved from imported cases to community transmission, and at the end of March 2020 Mexican authorities began to implement lockdowns and social distancing recommendations. As of November 2020, there had been 1.1 million confirmed COVID-19 cases and over 105,000 deaths.²⁰ The availability of data for each confirmed case of COVID-19 regarding outcome, age, sex, pre-existing health conditions, municipality of residence, and healthcare facilities where the patient received care has allowed for rich analyses of the risk factors associated with COVID-19 severity in Mexico. As in many other countries, the risk of more severe outcomes has been associated with non-communicable diseases, age, and male sex.²¹ Going further, some studies have already shown the association between socioeconomic disadvantage at the municipality level (analogous to the county level) and COVID-19 outcomes. Gutiérrez and Bertozzi²² and Millán-Guerrero et al.²³ found that higher poverty is positively associated with hospitalisation and death. Ortiz-Hernández and Pérez-Sastre²⁴ and Bello-Chavolla et al.²⁵ found a positive association between a social marginalisation index at the municipality level and the incidence of severe COVID-19 outcomes. Finally, Antonio-Villa et al.²⁶ explored the role of occupations, social lag, and population density at the municipality level on adverse COVID-19 outcomes, excess mortality, and vehicular mobility in Mexico City. They found that municipalities with high social lag indexes had the highest mortality risk.

In this paper we study COVID-19 hospitalisation and mortality in Mexico using a database from the Instituto Mexicano del Seguro Social (IMSS), one of the largest healthcare providers in Mexico. We use worker-level data from social security administrative records linked to patient records for people who were tested for SARS-CoV-2. Our focus is to test whether earnings are an important determinant of COVID-19 mortality, and if so, to analyse the possible factors behind the association. Unlike previous studies that analysed the effect of income at the patient level,^{18,19} our dataset includes information on pre-existing health conditions, and thus allows us to test whether comorbidities are a likely

mechanism behind the association between income and COVID-19 mortality. In contrast to the existing literature on the association of socioeconomic vulnerability and COVID-19 severity in Mexico,^{22–24,26} we are able to link COVID-19 outcomes to individual-level earnings data. Finally, lack of healthcare access is not a relevant variable here since all the individuals in our study had such access through the numerous facilities of IMSS (clinics and hospitals around the country).

Methods

Data sources

Mexico's public health system is highly segregated. People formally working in the private sector, the public sector, the armed forces, or the state-owned oil company are eligible to use health facilities that correspond to their type of employment. People who work in the informal sector, or shadow economy, do not have social security, but have access to limited health coverage through another subsystem, the Instituto de Salud para el Bienestar.^{27,28} Our data comes from the Mexican Social Security Institute (Instituto Mexicano del Seguro Social, IMSS), which is the agency in charge of providing social security benefits and medical services to all formally employed private sector workers and their families in Mexico. For this study we link data from two different types of administrative records. The first dataset includes information for 1.8 million people registered at IMSS who had COVID-19-like symptoms and went to a clinic for a diagnosis between March and November of 2020. Of this data, we use only that for patients who were tested for COVID-19 (56.2% of the total).

The health records of tested patients contain workers' social security numbers (SSN) and information from a triage questionnaire. The triage data includes information on self-reported pre-existing conditions such as asthma, cancer, chronic obstructive pulmonary disease, diabetes, HIV or other immunosuppressive conditions, hypertension and other cardiovascular diseases, obesity, kidney disease, smoking, and tuberculosis. Patients at IMSS were tested for COVID-19 with a polymerase chain reaction (PCR) test, from which we can identify those who tested positive for SARS-CoV-2. The records tell us whether patients were hospitalised for COVID-19, and whether they recovered from COVID-19 or died from it. The data covers COVID-19 outcomes through the middle of January 2021.

Using the SSN, we link the health records of tested patients to data on workers registered at IMSS and their wages. Besides the SSN, this dataset includes workers' daily earnings, sex, and age. Daily earnings are reported at their value on the last day of the month. We use data only for people who were employed and registered with IMSS between June 2019 and November 2020, who were at least 16 years old, and who were tested for COVID-19, a total of 422,053 registered and tested

workers. Finally, the sample is restricted to workers with a valid geographic indicator for their state of residence, which reduces the sample to 412,551 workers.

Outcomes

We estimate positivity percentages as the ratio of positive SARS-CoV-2 tests to the total number of positive and negative tests, multiplied by one hundred. We calculate the hospitalisation rate as the ratio of hospitalised individuals to the number of tests, multiplied by one hundred. Finally, we measure the case-fatality rate as the ratio of deaths to positive tests, multiplied by one hundred. We also examine the percentage of deaths among negative cases.

We analyse the outcomes by daily earnings centiles. Earnings are measured as the wage in the month of the test. If a worker is not employed during the month when they are tested, then we include the mean of non-zero earnings in the four months prior to being tested. If they did not work in those five months, then earnings are registered as zero. Most workers (96.6%) were employed at the time of the test or in the previous four months.

Data analysis

The statistical analysis is performed using Stata version 16. We first explore positivity and case-fatality rates with the data aggregated by daily earnings centiles. We also explore testing rates and positivity rates per 1000 workers by daily earnings centiles. We carry out a non-parametric estimate of the relationship between these aggregated rates and income percentiles.

For the COVID-19 outcomes, we estimate probit models using the linked patient-worker data (individual-level data) for each of the outcomes for Y_{it} : hospitalisation and death. The regression models are adjusted for earnings percentile (EarningsP_i), a dummy variable for a negative test (Negative), an interaction of earnings percentile and negative test, a vector of comorbidities (C_i), a vector of sex and age (D_i), testing-month fixed effects (μ_t), and state fixed effects (φ_s). This results in the following estimating equation:

$$\Pr(Y_{it}|X_{it}) = \Phi(\alpha + \beta \text{EarningsP}_i + \delta \text{Negative}_{it} + \theta \text{Negative}_{it} \times \text{EarningsP}_i + D_i \rho + C_i \gamma + \mu_t + \varphi_s)$$

where Y_{it} is an indicator variable equal to one if patient i was hospitalised or died, and $\Phi(\cdot)$ represents the normal cumulative distribution function. Given our observation that the positivity rate has a slightly negative relationship with daily earnings percentile, confirmed through regression analysis, we also adjust for the interaction between testing negative for SARS-CoV-2 and the earnings percentile. We estimate Eicker-White standard errors to correct for heteroskedasticity. To consider the difference between biological and chronological age²⁵, we perform an alternative estimation that replaces the vector of indicator variables of comorbidities

with the sum of comorbidities and the interactions of that number with sex and age. In addition to the estimations using the full sample, we estimate these models for subsamples of the population—women and men, age groups, and income levels—to explore whether the income gradient holds within groups (these results are available in the Supplementary Materials).

Results

The number of workers with a COVID test and a valid indicator for state of residence is 412,551, of whom 4.0% died. The case-fatality rate for those who tested positive for SARS-CoV-2 is 6.0% (see Table 1), in contrast to the 1.6% who tested negative. Workers who tested positive for SARS-CoV-2 were older, more likely to be male, more likely to have diabetes, and more likely to be obese or have hypertension than those who tested negative.

Fig 1 presents the hospitalisation rate and case-fatality rate by daily earnings centile, with markers for each ventile of the earnings distribution. We find that there is a negative monotonic relationship between each rate and the level of earnings. The probability of SARS-CoV-2 positive patients with the lowest incomes requiring hospitalisation is 25 percentage points (pp) greater than for the highest income patients, and their probability of dying is 13 pp greater. Fig 1 also shows a strong inverse association between these rates and earnings among people with negative SARS-CoV-2 tests.

The negative association shown in Fig 1 might be the result of selectivity bias in infection, testing, or positivity rates. For instance, it might be that poorer individuals have a higher probability of being infected, which should be reflected in higher testing or positivity rates. To gauge whether this is the case, Fig 2 plots the positivity and testing rates by earnings centile. These variables

	Tested	SARS-CoV-2 Negative	SARS-CoV-2 Positive	P-value
Observations	412,551	189,531 (46%)	223,020 (54%)	
COVID-19 Outcomes:				
Death	4%	1.6%	6%	0.000
Hospitalised	12.7%	7.9%	16.8%	0.000
Demographic variables:				
Age (years)	37.826	36.344	39.085	0.000
Female sex	50.1%	55.3%	45.7%	0.000
Proportion who reported comorbidities:				
Hypertension	12.4%	11.1%	13.6%	0.000
Other cardiovascular diseases	0.9%	0.9%	0.8%	0.000
Kidney disease	1%	1%	0.9%	0.262
Tuberculosis	0.3%	0.3%	0.3%	0.001
Cancer	0.2%	0.3%	0.2%	0.000
COPD	0.5%	0.5%	0.5%	0.018
Diabetes	8.3%	6.7%	9.6%	0.000
Asthma	3.1%	3.8%	2.6%	0.000
Other immunosuppression	0.9%	1%	0.8%	0.000
Smoking	7.7%	8.6%	7%	0.000
Obesity	15.8%	14.5%	16.9%	0.000
HIV positive	0.6%	0.7%	0.4%	0.000
Average daily earnings (\$MXN) by decile:				
Decile 1	127.5	128.0	127.1	0.000
Decile 2	153.2	153.4	153.0	0.000
Decile 3	196.4	196.5	196.3	0.180
Decile 4	237.9	237.5	238.1	0.000
Decile 5	293.7	293.5	293.8	0.127
Decile 6	365.0	364.7	365.2	0.042
Decile 7	468.6	469.2	468.1	0.001
Decile 8	595.9	595.9	595.8	0.797
Decile 9	778.3	777.8	778.7	0.169
Decile 10	1385.7	1396.0	1375.8	0.000

Table 1: Summary statistics of the estimating sample of tested workers registered at IMSS.

Notes: Authors' estimations using the linked employee-patient data for COVID-19. Column (1) shows all PCR-tested workers registered at IMSS. Columns (2) and (3) split the tested sample into negative and positive test results. Column (4) shows the p-value of the differences in means test for positive vs. negative.

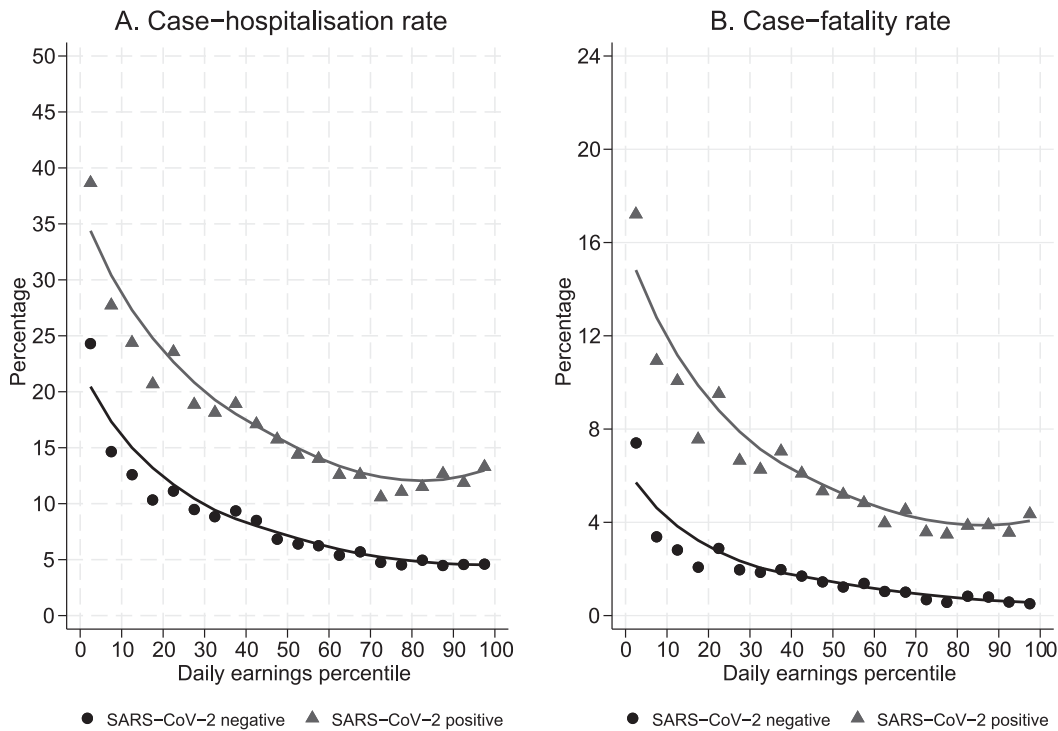


Figure 1. Hospitalisation and case-fatality rates by earnings. Notes: Authors’ estimations using linked data from workers registered with IMSS and patients tested for SARS-CoV-2 ($N = 422,053$). The figures show the percentage of hospitalised patients (Panel A) and fatalities (Panel B) per positive (grey) and negative (black) tests, by daily earnings centile. The solid lines are the non-parametric fit of the variable with daily earnings percentile. Each point represents a ventile of the earnings distribution.

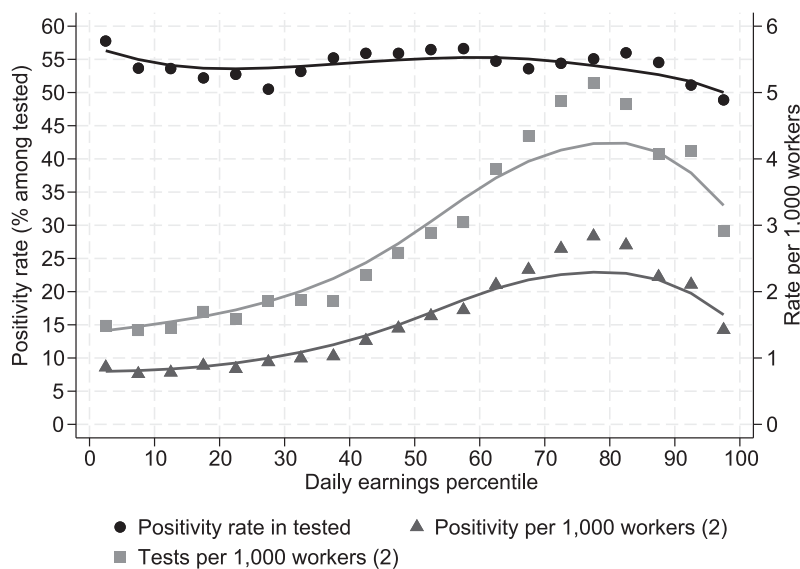


Figure 2. Positivity and testing rates by earnings. Notes: Authors’ estimations using linked data from workers registered with IMSS and patients tested for SARS-CoV-2. The positivity rate (confirmed cases/tested cases) is scaled on the left-hand axis ($N = 422,053$), and positivity and number of tests per 1000 workers (using the universe of workers) on the right-hand axis. Solid lines are the non-parametric fit regressions of the variables with daily earnings percentile. Each point represents a ventile of the earnings distribution.

do not show the same behaviour seen in hospitalisations and deaths by earnings in Fig 1, so selectivity bias cannot fully explain the income gradient: people with lower incomes did not get tested more than those with higher incomes and they did not have significantly higher positivity rates.

Table 2 presents the results of the estimations of the probability of hospitalisation or death among patients tested for SARS-CoV-2 (marginal effects). All the specifications control for age, female sex, month-of-year fixed effects, and state fixed effects. The first row of Columns (1) and (4) shows the association of the income percentile with the probability of being hospitalised and the probability of dying among those who tested positive without adjusting for comorbidities. We find that moving from the poorest to the richest (from centile 0 to 1) is associated with a decrease in the probability of hospitalisation of 0.146 (95% CI: -0.150, -0.142) and a decrease in the probability of dying of 0.027 (95% CI: -0.028, -0.025). Columns (2) and (5) show the comorbidity-adjusted association of the two outcomes with the income percentile. As expected, those who tested negative for SARS-Cov-2 have a three-percentage point lower probability of being hospitalised and a one percentage point lower probability of dying. In addition, given that the interaction of the earnings percentile and the negative test indicator is negative, we also find a less steep income gradient among those who tested negative than among those who tested positive, which confirms the finding in Fig 1.

Table 2 also shows that most of the comorbidities have the expected significant effects that have been identified in other studies. We control for comorbidities in two different ways. Columns (2) and (4) control for indicator variables of each comorbidity. In turn, Columns (3) and (6) control for a patient's total number of comorbidities and the interaction of this number with age and female sex. Adjusting for comorbidities in either way does not significantly change the income gradients for hospitalisation or death.

Fig 3 shows the predicted probability of hospitalisation (Panel A) and the predicted probability of dying (Panel B) for the models in Columns (2) and (5) of Table 2, respectively, evaluated for each decile and using mean values for the rest of the independent variables. Among those who tested positive, the poorest workers have a probability of hospitalisation approximately four times greater than the richest workers, and a probability of dying that is five times greater. These differences are obtained while controlling for different types of comorbidities, and it should be emphasised that all persons in this study had access to healthcare. Fig 3 also shows that the probability of dying or of being hospitalised is greater for the poorest workers who tested negative than for those in the upper part of the wage distribution who tested positive.

Discussion

To date, the literature has pointed to associations of COVID-19 mortality with (at most) area-level socioeconomic measures, such as poverty or median income. Analyses based on aggregated data may lead to an ecological fallacy: grouped data can hide within-group heterogeneity that reveals different associations between the variables of interest. Two notable studies that used individual-level data on COVID-19 outcomes and income found evidence of a negative income gradient, but they were not able to adjust their estimates for individual-level comorbidities, and they also leave open the question of whether people with low income are at higher risk of infection.^{18,19}

In this paper, we present evidence that suggests that neither the positivity rate nor the testing rate by income level explain the income gradient in COVID-19 mortality. We also show that among people who tested positive, those with lower incomes had, in fact, four times the probability of being hospitalised, which means that they experienced more severe disease than richer people. Those in the lowest income decile also had five times the probability of dying as those in the highest decile. In comparison, Decoster et al. found that those in the lowest income decile in Belgium are twice as likely to die than those in the highest decile.¹⁹ Drefahl et al. found that those men in the lowest tertile in Sweden are 75% more likely to die from COVID-19 than those in the highest tertile.¹⁸ To further place our estimates in context, among those who tested positive (Table 2, Column 6), the difference in the probability of dying between the richest and the poorest in Mexico is approximately equal to the average marginal effect of being a man or of being immunosuppressed, or twice the effect of having diabetes, a leading risk factor in the Mexican population.²¹ In Mexico, none of the previous studies of the association of socioeconomic status with COVID-19 outcomes estimated odds ratios above two when comparing the most and the least vulnerable.^{23,24,26} The use of individual-level data allows us to uncover greater inequalities than previously found in the impact of the pandemic with respect to socioeconomic status.

As shown in the Supplementary Materials, several of the pre-existing conditions associated with higher COVID-19 mortality also exhibit a negative income gradient. Hence, a higher rate of comorbidities could be the underlying cause behind the negative income gradient in COVID-19 mortality. In this paper, we show that even after adjusting our estimations for individual-level comorbidities, the negative income gradient persists in hospitalisation and fatality rates owing to COVID-19. There is thus an association of individual income with COVID-19 hospitalisations and mortality that is independent of comorbid diseases.

Another advantage of our data is that, by definition, all the people in our sample have access to healthcare,

	Pr (Hospitalisation X)			Pr(Death X)		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily earnings percentile	-0.1459*** [0.0019]	-0.1306*** [0.0019]	-0.1371*** [0.0019]	-0.0266*** [0.0007]	-0.0215*** [0.0006]	-0.0217*** [0.0006]
SARS-CoV-2 negative	-0.0301*** [0.0017]	-0.0336*** [0.0017]	-0.0322*** [0.0016]	-0.0109*** [0.0006]	-0.0111*** [0.0005]	-0.0101*** [0.0005]
Earnings percentile x SARS-CoV-2 negative	-0.0449*** [0.0031]	-0.0382*** [0.0031]	-0.0415*** [0.0031]	-0.0118*** [0.0011]	-0.0092*** [0.0010]	-0.0097*** [0.0010]
Comorbid diseases controls:						
Diabetes		0.0656*** [0.0020]			0.0107*** [0.0006]	
Obesity		0.0146*** [0.0012]			0.0072*** [0.0004]	
Hypertension		0.0127*** [0.0013]			0.0033*** [0.0004]	
Other cardiovascular disease		0.0418*** [0.0053]			0.0007 [0.0011]	
Kidney disease		0.2863*** [0.0094]			0.0591*** [0.0037]	
Tuberculosis		0.0289*** [0.0087]			0.0098*** [0.0031]	
Cancer		0.0495*** [0.0116]			0.0087*** [0.0035]	
Chronic obstructive pulmonary disease (COPD)		0.0233*** [0.0062]			0.0026* [0.0015]	
Asthma		-0.0097*** [0.0025]			-0.0027*** [0.0007]	
Smoking		-0.0089*** [0.0015]			-0.0019*** [0.0004]	
HIV positive		0.0015 [0.0057]			0.0069*** [0.0024]	
Other immunosuppression		0.0965*** [0.0070]			0.0236*** [0.0027]	
Number of comorbidities (#Comorb)			0.0258*** [0.0023]			0.0126*** [0.0006]
#Comorb x female sex			0.0281*** [0.0036]			0.0022** [0.0010]
#Comorb x age			0.0001 [0.0001]			-0.0002*** [0.0000]
#Comorb x female sex x age			-0.0008*** [0.0001]			-0.0000 [0.0000]
Additional controls:						
Age	0.0060*** [0.0000]	0.0052*** [0.0000]	0.0055*** [0.0000]	0.0015*** [0.0000]	0.0013*** [0.0000]	0.0014*** [0.0000]
Female sex	-0.0764*** [0.0009]	-0.0736*** [0.0009]	-0.0699*** [0.0011]	-0.0201*** [0.0004]	-0.0184*** [0.0004]	-0.0187*** [0.0004]
Month of year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	412,551	412,551	412,551	412,551	412,551	412,551
Mean among SARS-CoV-2 positive	0.168	0.168	0.168	0.060	0.060	0.060

Table 2: Estimations of the probability of hospitalisation and of death.

Notes: Each column presents the estimation of a probit regression (marginal effects). Columns 1-3 are estimates of the probability of tested patients being hospitalised; columns 4-6 are estimates of their probability of dying. The three models of each outcome control for a different set of control variables. The first model controls for earnings percentile, testing negative for SARS-CoV-2, and the interaction of testing negative and earnings percentile. The second model adds indicator variables for comorbidities. The third model adds to the first the number of comorbidities a worker has and the interactions between that number, female sex, and age. All regressions adjust for age, female sex, month-of-year indicator variables and state fixed effects. Robust standard errors in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

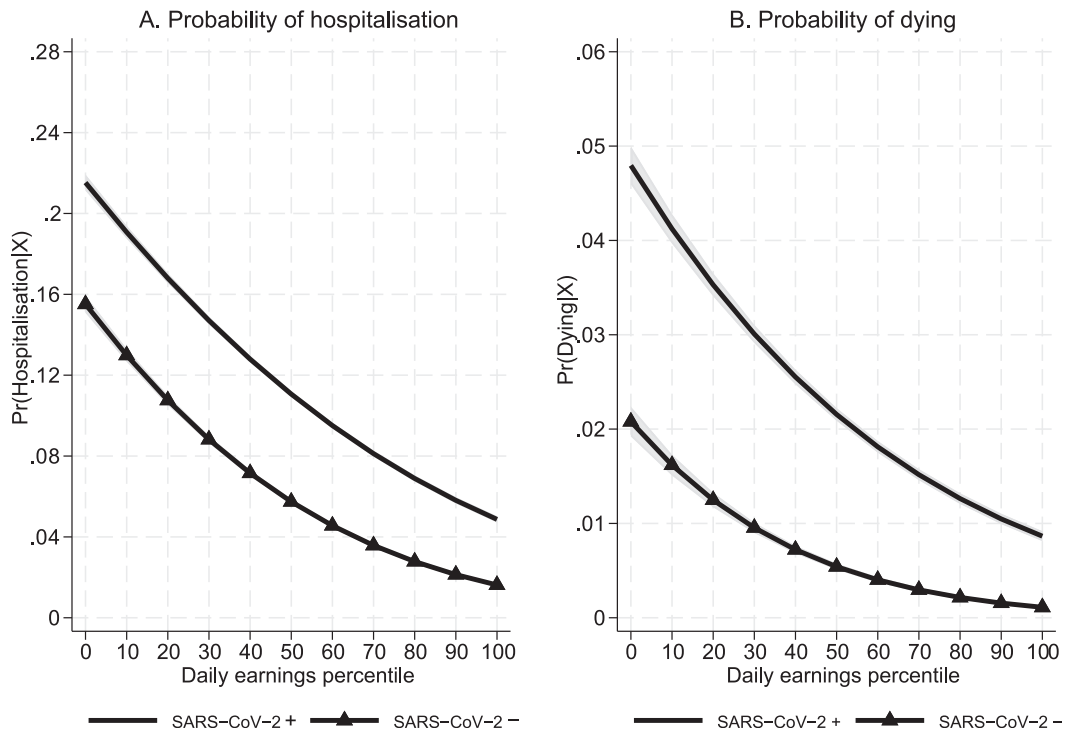


Figure 3. Income gradient in the probability of hospitalisation and dying. Notes: Authors' estimations using linked data from workers registered with IMSS and patients tested for SARS-CoV-2. The panels present the predicted probability of hospitalisation (Panel A), and the predicted probability of dying (Panel B) for the models in Columns (2) and (5) of Table 2, respectively, evaluated at each decile and using mean values for the rest of the independent variables for those who tested positive or negative. Confidence intervals are at the 95% level.

and in fact they were all tested for SARS-CoV-2 at a public health clinic. Any observed difference in either the positivity or the case fatality rate is therefore not explained by lack of access to healthcare, as suggested in previous literature³, but could still be due to disparities in obtaining timely and adequate care, which we are not able to control for in our analysis. Another possible explanation for the income gradient is that lower-income people may wait longer to get healthcare, so that when they finally go to the hospital, their illness is already severe. These are issues that we leave for future exploration. Notwithstanding these possibilities, our results provide support for the characterisation of COVID-19 in Mexico as a syndemic,^{8–10} and not just a pandemic.

Our data has some limitations. It is only representative of formal workers registered at IMSS who were suspected of having COVID-19 and were tested. Therefore, our conclusions do not apply to the Mexican population overall: our data do not include workers registered with other health systems, workers in the informal sector, or people without employment. These different employment statuses might themselves involve different profiles and severity of comorbid diseases, and different income levels. Informal sector workers, for example,

who represent more than half of the working population in Mexico, have lower wages and less access to healthcare than those working in the formal sector. Our study sample thus has higher incomes and less inequality, and a lower incidence of comorbid diseases, than the population at large. This bias, however, would tend to reduce an income gradient, so it is likely that the income gradient for the rest of the Mexican population is even steeper. In addition, we cannot control for actual obesity, as measured by body mass index, or for ethnic origin, as these variables are not observed in the data. Overall, our findings fill a gap in the literature by showing that the income gradients in COVID-19 hospitalisation and mortality rates are independent of comorbidities and of access to healthcare among workers registered at IMSS.

Role of the funding source

This study was not supported by any external funding source.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgments

We thank the three anonymous reviewers for their excellent comments and suggestions that substantially improved the paper. We are grateful for superb research assistance from Raquel Yunoen Badillo. The content of this study does not necessarily reflect the opinion of the Banco de Mexico or its Board, or the Instituto Mexicano del Seguro Social (IMSS). Responsibility for the information, errors, omissions, and views expressed here lies entirely with the authors.

Data sharing statement

Access to the dataset used in this study is restricted to IMSS employees. IMSS affiliated authors had access to the full dataset and all authors had access to de-identified data. Full dataset is not available due to legal restrictions (Ley del Seguro Social, Article 22). Sharing of de-identified aggregated data will be considered by IMSS authors upon reasonable request.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:[10.1016/j.lana.2021.100115](https://doi.org/10.1016/j.lana.2021.100115).

References

- Pan D, Sze S, Minhas JS, Bangash MN, Pareek N, Divall P, et al. The impact of ethnicity on clinical outcomes in COVID-19: a systematic review. *EClinicalMedicine* 2020;23:100404.
- Upshaw TL, Brown C, Smith R, Perri M, Ziegler C, Pinto AD, Camacho-Rivera M. Social determinants of COVID-19 incidence and outcomes: a rapid review editor *PLoS One* 2021;16(3):e0248336.
- Abrams EM, Szefer SJ. COVID-19 and the impact of social determinants of health. *Lancet Respir Med* 2020;8:659–61.
- Yancy CW. COVID-19 and African Americans. *JAMA* 2020;323:1891–2.
- Oronce CIA, Scannell CA, Kawachi I, Tsugawa Y. Association between state-level income inequality and COVID-19 cases and mortality in the USA. *J Gen Intern Med* 2020;35:2791–3.
- Gu T, Mack JA, Salvatore M, Prabhu Sankar S, Valley TS, Singh K, et al. Characteristics Associated with racial/ethnic disparities in COVID-19 outcomes in an academic health care system. *JAMA Netw Open* 2020;3(10):e2025197.
- Fielding-Miller RK, Sundaram ME, Brouwer K, Zaller ND. Social determinants of COVID-19 mortality at the county level editor *PLoS One* 2020;15(10):e0240151.
- Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. *J Epidemiol Community Heal* 2020;74(11):964–8.
- Horton R. Offline: COVID-19 is not a pandemic. *Lancet* 2020;396(10255):874.
- Mendenhall E. The COVID-19 syndemic is not global: context matters. *Lancet* 2020;396(10264):1731.
- 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.
- Little C, Alsen M, Barlow J, Naymagon L, Tremblay D, Genden E, et al. The impact of socioeconomic status on the clinical outcomes of COVID-19: a retrospective cohort study. *J Community Health* 2021: 1–9.
- Abedi V, Olulana O, Avula V, Chaudhary D, Khan A, Shahjouei S, et al. Racial, economic, and health inequality and COVID-19 infection in the United States. *J Racial Ethn Heal Dispar* 2020: 1–11.
- Brandily P, Brébion C, Briole S, Khoury L. A poorly understood disease? The unequal distribution of excess mortality due to COVID-19 across French municipalities. medRxiv; 2020. <https://doi.org/10.1101/2020.07.09.20149955> [cited 2021 Mar 30].
- Chaudhry R, Dranitsaris G, Mubashir T, Bartoszko J, Riazi S. A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine* 2020;25:100464.
- Godefroy R, Lewis J, Alsan M, Bobonis G, Brodeur A, Lange F. What explains the socioeconomic status-health gradient? Evidence from workplace COVID-19 infections. medRxiv; 2021. <https://doi.org/10.1101/2021.03.23.21254170> Mar 24 [cited 2021 Apr 5].
- Sundaram ME, Calzavara A, Mishra S, Kustra R, Chan AK, Hamilton MA, et al. The individual and social determinants of COVID-19 in Ontario, Canada: a population-wide study. medRxiv; 2020. <https://doi.org/10.1101/2020.11.09.20223792> [cited 2021 Apr 5].
- Drefahl S, Wallace M, Mussino E, Aradhya S, Kolk M, Brandén M, et al. A population-based cohort study of socio-demographic risk factors for COVID-19 deaths in Sweden. *Nat Commun* 2020;11(1):1–7.
- Decoster A, Minten T, Spinnewijn J. The Income Gradient in Mortality during the Covid-19 Crisis: Evidence from Belgium. *J Econ Inequal* 2021;19:551–570. <https://doi.org/10.1007/s10888-021-09505-7>
- Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Giattino C, Ortiz-Ospina E, et al. Coronavirus Pandemic (COVID-19). Published online at OurWorldInData.org. Retrieved from: <https://ourworldindata.org/coronavirus>. [Online Resource] [cited 2021 Aug 20].
- Bello-Chavolla OY, Bahena-López JP, NE AV, Vargas-Vázquez A, González-Díaz A, Márquez-Salinas A, et al. Predicting mortality due to SARS-CoV-2: a mechanistic score relating obesity and diabetes to COVID-19 outcomes in Mexico. *J Clin Endocrinol Metab* 2020;105(8):2752–61.
- Gutierrez JP, Bertozzi SM. Non-communicable diseases and inequalities increase risk of death among COVID-19 patients in Mexico. *PLoS One* 2020;15(10):e0240394.
- Millán-Guerrero RO, Caballero-Hoyos R, Monárrez-Espino J. Poverty and survival from COVID-19 in Mexico. *J Public Health (Oxf)* 2020: 1–8.
- Ortiz-Hernández L, Pérez-Sastré MA. Social inequalities in the progression of COVID-19 in the Mexican population. *Rev Panam Salud Pública* 2020;44:25e106–e106. <https://doi.org/10.26633/RPSP.2020.106>.
- Bello-Chavolla OY, González-Díaz A, NE AV, Fermín-Martínez CA, Márquez-Salinas A, Vargas-Vázquez A, et al. Unequal impact of structural health determinants and comorbidity on COVID-19 severity and lethality in older Mexican adults: considerations beyond chronological aging. *J Gerontol Ser A* 2021;76(3):25e52–9.
- Antonio-Villa NE, Fernandez-Chirino BSci L, Pisanty-Alatorre J, Mancilla-Galindo J, Kammar-García A, Vargas-Vázquez A, et al. Comprehensive evaluation of the impact of sociodemographic inequalities on adverse outcomes and excess mortality during the COVID-19 pandemic in Mexico City. medRxiv; 2021 Mar 12 [cited 2021 Apr 6]; Available from <https://doi.org/10.1101/2021.03.11.21253402>.
- Knaut FM, González-Pier E, Gómez-Dantés O, García-Junco D, Arreola-Ornelas H, Barraza-Lloréns M, et al. The quest for universal health coverage: achieving social protection for all in Mexico. *Lancet* 2012;380(9849):1259–79.
- Reich MR. Restructuring health reform, Mexican style. *Health Syst Reform* 2020;6(1):1–11.